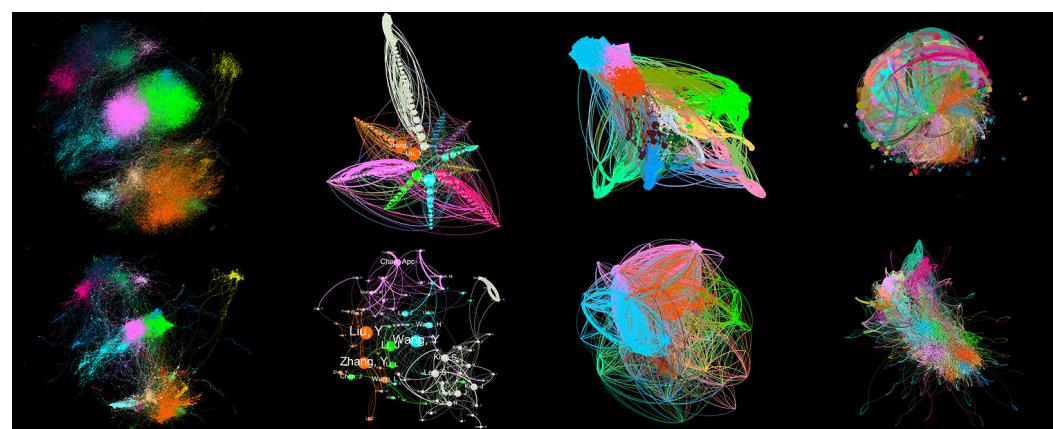
Visualizzare la conoscenza. La rappresentazione delle reti citazionali internazionali nell'ambito delle scienze grafiche

Amedeo Ganciu
Andrea Sias

Abstract

Questo articolo sperimenta la visualizzazione di big data relativi alle reti citazionali all'interno degli ambiti di ricerca di interesse delle scienze grafiche. A partire da precedenti ricerche svolte negli ultimi anni anche in occasione dei convegni UID da docenti afferenti al settore ICAR/17, attraverso una massiva ricerca all'interno del database Web of Science, è stato possibile individuare e analizzare numericamente e graficamente oltre 100'000 documenti per ricostruire le reti di citazione tra gli articoli, le collaborazioni tra gli autori e le reti di citazioni tra le keywords presenti su ciascuna pubblicazione. In questo articolo si sperimenta in particolare l'applicazione degli strumenti della Social Network Analysis per l'indagine e la mappatura del dialogo tra gli ambiti del sapere e i ricercatori che li rappresentano. Nella consapevolezza di alcuni limiti connaturati alla natura della ricerca, questo studio si configura come una prima sperimentazione che applica metodi e strumenti consolidati nella letteratura internazionale ad un caso di studio inedito. La ricerca non può dirsi conclusa e completa, in quanto riferita esclusivamente a dati di tipo bibliometrico provenienti da un unico database (*WoS*) che necessiterebbero di essere integrati con i dati provenienti da altri database (*Scopus*) e dai dati relativi ai prodotti di ricerca non bibliometrici che attualmente risultano essere prevalenti all'interno della comunità scientifica ICAR/17.

Parole chiave
Data visualization, big data, grafi, reti citazionali, scienze grafiche

Topic
Mappare

Elaborazione grafica delle reti. (elaborazione Amedeo Ganciu).

doi.org/10.3280/oa-832-c155

Introduzione

La ricerca inter e transdisciplinare sta suscitando molto interesse tra i ricercatori per molti motivi e con differenti risultati [Cicalò 2020a], non solo perché consente una visione unitaria nella quale i confini disciplinari sono assunti come riferimento e non come limite [Luigini, 2021a], ma anche per il potenziale economico associato ad una contaminazione e trasferimento di saperi tra le discipline [Jin et al. 2011]. Tra le tecniche di rappresentazione della complessità dei *big data,* diversi ricercatori stanno evidenziando la centralità dell'infografica e della visualizzazione dati [Bocconcino e Vozzola, 2020, Cicalò e Menchetelli, 2020; Gigliarelli et al., 2020]. In particolare, tra gli strumenti di supporto alla conoscenza dei sistemi dinamici complessi, la modellazione indirizzata al paradigma reticolare sta acquisendo una crescente attenzione da parte della comunità scientifica internazionale con numerosi esempi applicativi [Newman and Girvan 2004; Boccaletti et al. 2006], nei sistemi economici e sociali [Robins 2013; Ganciu et al. 2016; Green 2007], infrastrutture di trasporto e della mobilità [Ganciu et al. 2018; Porta et al. 2006; Xie and Levinson 2011; Derrible and Keendy, 2011], pianificazione urbana e territoriale [Cook 2002; Estrada and Bodin 2008; Urban, 2009], fornendo gli strumenti per una migliore comprensione del fenomeno [Balestrieri and Ganciu 2019]. Newman (2003) identifica quattro categorie di reti: reti sociali, come le forme di contatto o interazione tra gli individui o le reti di citazioni accademiche; reti di informazione, come i collegamenti nel World Wide Web; reti tecnologiche, come i sistemi idrici, di trasporto e di energia; e reti biologiche, come le reti alimentari con prede e predatori. In tutti e quattro i casi, la base metodologica per l'analisi e rappresentazione delle reti può essere fatta risalire allo sviluppo della teoria dei grafi da parte di Eulero nel 1736, che per primo pose le basi della disciplina indicata come Teoria dei Grafi nella risoluzione del problema dei Ponti di Konigsberg.(fig. 01).

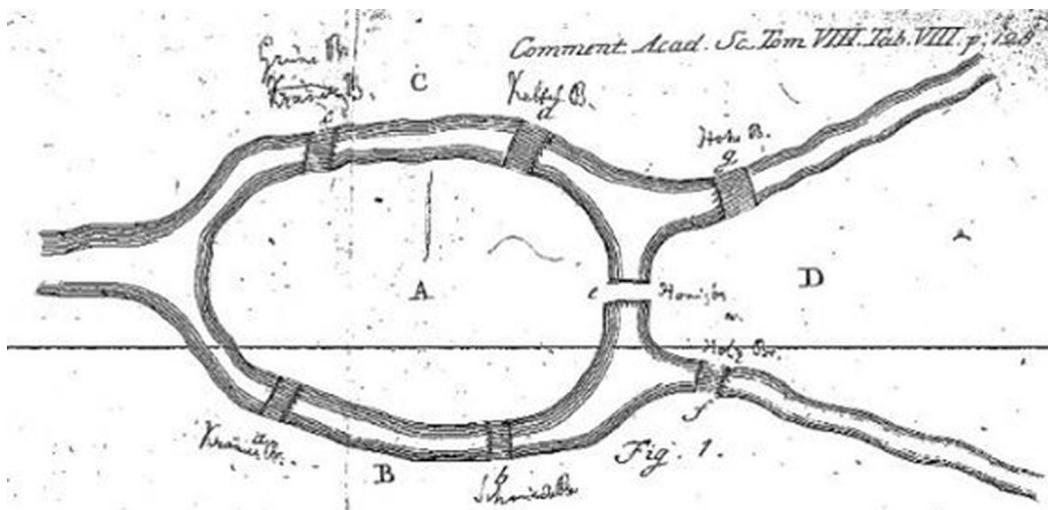



Fig. 01. Mappa della cittadina di Konigsberg realizzata da Eulero per la risoluzione del problema.

All'interno della categoria delle reti sociali, le reti bibliografiche rappresentano un esempio di rete non spazializzata, in cui la posizione euclidea di nodi e archi perde di significato a discapito dell'importanza assunta dai nodi attraverso le relazioni che questi stabiliscono tra loro, ossia del numero di citazioni ricevute. Infatti, il numero di citazioni che riceve un documento da altri documenti è il fattore che determina la sua permanenza o la sua transitorietà nella finestra di visibilità di una rete citazionale, come già dimostrato fin dai lavori pionieristici degli anni '70 [Garfield et al. 1964; Batagelj 2003; Börner et al. 2003; Luigini 2021b] e ancora oggi, all'interno del cyberspazio scientifico, è possibile identificare strutture di citazioni ricorrenti basate su lavori con un alto numero di citazioni, circondate da altri lavori, che tendono ad essere marginali e ad avere una natura transitoria nel panorama scientifico [Abt 1998].

Nel particolare di questa ricerca, precedenti studi hanno già dimostrato come il settore del Disegno sia caratterizzato da un elevata eterogeneità sul piano nazionale [Luigini 2021a e 2021b], in cui confluiscono contributi delle *hard* e delle *humanities sciences*.
Utilizzando questi lavori come spunto iniziale, si è proceduto alla ricerca all'interno del database Web of Science dei documenti e degli articoli riconducibili alle scienze grafiche [Cicalò 2020b]; in particolare è stato possibile individuare e analizzare numericamente e graficamente oltre 100'000 documenti che hanno permesso di ricostruire le reti di citazione tra gli articoli, le collaborazioni tra gli autori e le reti di citazione tra le keywords presenti su ciascuna pubblicazione.
Attraverso l'analisi delle citazioni tra articoli si è ricostruita la nuvola di pubblicazioni individuando i lavori più significativi, cioè quelli con maggiore numero di citazioni e che potenzialmente hanno le maggiori probabilità di influenzare l'indirizzo delle successive ricerche.
Ricostruendo la rete degli autori è possibile misurare e visualizzare come gli studiosi collaborano tra loro e con quale intensità; infine, con la rete di citazione delle keywords delle ricerche è possibile osservare quali sono le principali tematiche e come sono connesse tra loro. La ricerca sperimenta inoltre l'applicazione di differenti algoritmi per la visualizzazione e il layout automatizzato delle reti, assolutamente necessari considerando la complessità delle reti in termini di nodi e archi, ottenendo risultati differenti per semplicità di lettura e interpretazione dei risultati.
Questa ricerca non ha la pretesa di considerarsi esaustiva dell'intera disciplina per diversi limiti, primo tra tutti la non afferenza bibliometrica del settore italiano; se infatti diverse riviste di settore sono indicizzate da sistemi come Web of Science o Scopus [Luigini 2021b], un significativo numero di lavori è stato pubblicato su riviste non censite sul profilo bibliometrico e quindi sfuggono a questa analisi. Per questo motivo l'indagine si deve considerare come un primo elemento verso una futura ricerca più integrale e maggiormente focalizzata sul panorama nazionale che mira ad evidenziare il ruolo italiano all'interno del dibattito internazionale.

Metodo e Dati

I dati per l'analisi sono stati prelevati dal database online *Web of Science* (WoS) attualmente gestito dalla *Clarivate Analytics*, che indicizza i lavori pubblicati su oltre 12.000 riviste e 160.000 atti di conferenze, per un totale di oltre 90 milioni di articoli e oltre un miliardo di riferimenti citati [Reuters 2016].
La selezione degli articoli è stata eseguita cercando nel loro titolo la presenza di parole ritenute significative all'interno del settore disciplinare e come già detto utilizzate in precedenti ricerche come da Luigini per il 42° Convegno Internazionale dei Docenti delle Discipline della Rappresentazione - Congresso della Unione Italiana per il Disegno (tab.1).
La parola "Disegno" anche se non espressamente indicata nella precedente ricerca è stata volutamente aggiunta considerato il suo uso anche all'interno della letteratura in lingua inglese. Successivamente, per ciascuna voce si è proceduto ad un'ulteriore selezione e raffinazione utilizzando le categorie disciplinari indicate all'interno del WoS maggiormente attinenti al settore disciplinare delle scienze grafiche (tab. 2).
Quindi, attraverso la query incrociata: Select = ["Parola nel Titolo" AND "Categoria WoS"] è stato possibile individuare inizialmente un totale di 260'002 record.

Parola titolo	N°	Parola titolo	N°	Parola titolo	N°
Modeling	72753	Survey	3813	Integrated Survey	63
Design	59567	Knowledge	3112	HBIM	58
City	28266	Representation	2362	Digital Survey	48

Architecture	22218	Communication	2053	Disegno	32
Project	18242	BIM	1920		
History	11211	Cultural Heritage	1291		
Landascape	10219	3D Model	1191		
Image	7630	Sign	1092		
Drawing	6154	Virtual Reality	449		
Perspective	6010	Augmented Reality	248	Totale	260002

Tab. 1. Parole utilizzate per cercare le pubblicazioni scientifiche all'interno di WoS.

Categorie disciplinari selezionate all'interno di Web of Science	
Art	Urban Studies
Architecture	Regional Urban Planning
Engineering Civil	

Tab. 2. Categorie disciplinari di WoS utilizzate per filtrare le pubblicazioni.

Con lo scopo di evitare di analizzare record duplicati, giacché all'interno dello stesso titolo possono comparire due, tre o più parole contemporaneamente, comportando un duplice, triplice etc, conteggio dell'informazione, la base di analisi è stata raffinata identificando in maniera univoca ciascuna pubblicazione, con eliminazione dei record duplicati. Questa delicatissima e fondamentale procedura atta a garantire la correttezza di ogni successiva analisi è stata eseguita attraverso la lettura automatizzata del codice DOI (*Digital Object Identifiers*) univoco per ciascun record presente all'interno del WoS, consentendo di ottenere un database "pulito" di 101'741 record, ossia le pubblicazioni utilizzate in questa ricerca. Per ciascuno dei record selezionati, si è scaricato dal WoS il relativo ISI.file, che è una sorta di scheda metadati, contenente tutte le informazioni principali del documento come gli autori, l'abstract, le keywords, la rivista o libro su cui è stato pubblicato e, aspetto fondamentale per le reti di citazione, i documenti che cita tra le sue voci di bibliografia; ciascuna di queste informazioni è classificata all'interno del ISI.file attraverso dei TAG che i sistemi bibliometrici riconoscono e possono utilizzare. La base dati complessiva, composta quindi da 101'741 ISI.file, è stata pre-processata all'interno della piattaforma Sci2Tool, realizzata dalla *Cyberinfrastructure for the Network Science Center* - Indiana University (USA). Nel particolare si sono modellate le reti degli "Autori", delle "Keywords" e delle "Citazioni tra pubblicazioni", differenti per finalità di ricerca, complessità in termini di numero di nodi e archi, e tipologia in termini di orientamento e peso degli archi (tab. 3).

Rete	**Orientamento della rete**	**Peso degli Archi**
Autori	Non Orientata	Si
Keywords	Non Orientata	Si
Citazioni tra articoli	Orientata	No

Tab. 3. Reti analizzate con indicazione della tipologia

Una rete si può rappresentare attraverso la sua formalizzazione matematica; attraverso la sua rappresentazione grafica; oppure attraverso la sua forma di matriciale (fig. 02). Nel merito della disciplina matematica e delle sue applicazioni, con gli attuali algoritmi è possibile calcolare l'influenza reciproca tra i diversi nodi della rete attraverso la computazione degli archi con peso e direzione diversi [Mathis 2007]. Per esempio, un grafo si definisce orientato se ci sono obblighi o regole che determinano la direzione o percorribilità di un arco (come nella strada a senso unico), un grafo può essere pesato se è possibile assegnare un'importanza diversa a ciascun arco, come espressione dell'intensità delle relazioni tra i nodi [Caldarelli 2007].

	e1	e2	e3	e4	e5
v1	1	0	0	1	0
v2	1	1	0	0	1
v3	0	1	1	1	0
v4	0	0	1	0	1

Fig. 02. Esempio delle tre rappresentazioni dello stesso grafo. Elaborazione Amedeo Ganciu.

Nella rete degli Autori ciascun nodo rappresenta un autore e l'esistenza di un arco tra due nodi indica una loro collaborazione su una pubblicazione, al crescere della frequenza con cui questo evento si registra crescerà anche il peso degli archi; quindi, la forza che in qualche modo li lega; la rete non è orientata giacché affermare che "A" collabora con "B" è uguale a dire che "B" collabora con "A". Con la rete delle "Keywords" è possibile estrapolare e rappresentare le informazioni su quali siano gli argomenti di ricerca che attraggono maggiormente i ricercatori e come siano correlati tra loro attraverso le relazioni di citazione tra articoli; anche questa rete non è orientata, perché come nel caso precedente non avrebbe significato una direzionalità, mentre ha significato l'attribuzione di un peso agli archi, indicatore della frequenza con cui due keyword sono connesse da una relazione di citazione.
La rete delle "Citazioni tra articoli" indica ovviamente le relazioni tra le pubblicazioni ossia "chi cita chi", tutti gli archi hanno importanza uguale quindi la rete non è pesata, ma al contrario dei casi precedenti, è orientata, perché anche intuitivamente è logico pensare che se l'articolo "A" cita l'articolo "B", non è possibile che accada contemporaneamente anche il contrario. Successivamente alla costruzione delle tre reti, si sono svolte le basilari analisi quantitative sulle strutture, avvalendosi del software opensource *Gephi*.
In particolare, per avere prime sommarie informazioni sulla loro morfologia si è eseguito per ciascuna rete il calcolo del grado del nodo (Deg), ossia la somma del numero di archi che convergono su ciascun nodo della rete [Foulds 2012] e nel caso delle reti pesate la somma di tutti i pesi di tutti gli archi che convergono su ciascun nodo della rete. Inoltre, in tutti e tre i casi si è eseguita l'analisi per individuare le Community attraverso l'applicazione dell'algoritmo di *Louviain* [Blondel et al. 2008].
Quest'ultima analisi permette di determinare la presenza dei cluster all'interno di una rete, caratterizzati per avere al loro interno la massima densità di relazioni (archi) e contemporaneamente minima tra nodi appartenenti a Community differenti [De Meo et al. 2011]. In ultimo, si sono applicati differenti algoritmi per la visualizzazione delle reti appartenenti alla tipologia che enfatizzano le divisioni, le complementarità e il ranking tra i nodi che compongono le reti. Nello specifico per massimare a livello visivo le sub-strutture o Community che compongono una rete si è utilizzato l'algoritmo "OpenOrd" [Martin et al. 2011], che si aspetta una rete non orientata e pesata e può elaborare rappresentazioni di network con un numero di nodi compreso tra 100 e 1'000'0000.
Tra gli algoritmi di visualizzazione delle complementarità, si sono scelti il "*ForceAtlas2*", nativo all'interno di *Gephi* e sviluppato da Mathieu Jacomy, e lo Yifan Hu Multilevel layout sviluppato nel 2005 [Hu 2005].
Entrambi gli algoritmi gestiscono la visualizzazione di reti con un massimo di 1'000'000 di nodi, però il ForceAtlas2 permette scegliere se usare il peso degli archi per ottenere una migliore visualizzazione delle reti, cosa che invece non computa lo Yifan Hu Multilevel layout. Infine, tra gli algoritmi che enfatizzano il ranking si è scelto di utilizzare il *Radial Axis Layout*, nativo all'interno di *Gephi* e sviluppato da Matt Groeninger, che consente di creare ordinamenti per elementi caratterizzati da proprietà omofile come l'appartenenza alla stessa community o altra caratteristica, anche lui è capace di gestire reti con complessità fino ad 1'000'000 di nodi (fig. 03).

Risultati

La prima fase di screening eseguita all'interno del portale *WoS* indica alcune delle parole cercate all'interno dei titoli delle come significativamente preponderanti in termini di numerosità statistica, per esempio "Modeling" e "Design" presenti in diverse decine di migliaia di pubblicazioni, mentre altri termini come "Disegno" e "Digital Survey" significativamente più marginali rispetto alle prime. Considerando le categorie disciplinari utilizzate all'interno di *WoS* per indicizzare le pubblicazioni si può osservare come la classe afferente all'ingegneria civile e architettura sia preponderante per la maggior parte delle ricerche, in modo particolare per quelle che contengono nel titolo i termini "Modeling", "Design", "Project" e "Architecture". Le classi della pianificazione territoriale e urbanistica inglobano invece la maggior parte dei documenti che contengono le parole "*Perspective*" e "Knowledge". Infine, nella maggior parte dei casi i titoli che contengono le parole "Drawing" e "Image" sono classificati da *WoS* come afferenti alla classe Arte. (fig. 04).
L'estrapolazione delle tre reti dagli ISI.file ha generato strutture significativamente differenti per complessità in termini di numerosità dei nodi e connessioni, ma abbastanza simili nella morfologia distributiva, costituite da un differente numero di Community che però presentano una qualità nella loro separazione espressa attraverso il valore della *Modularity* molto elevato per tutti e tre i casi (tab.4).

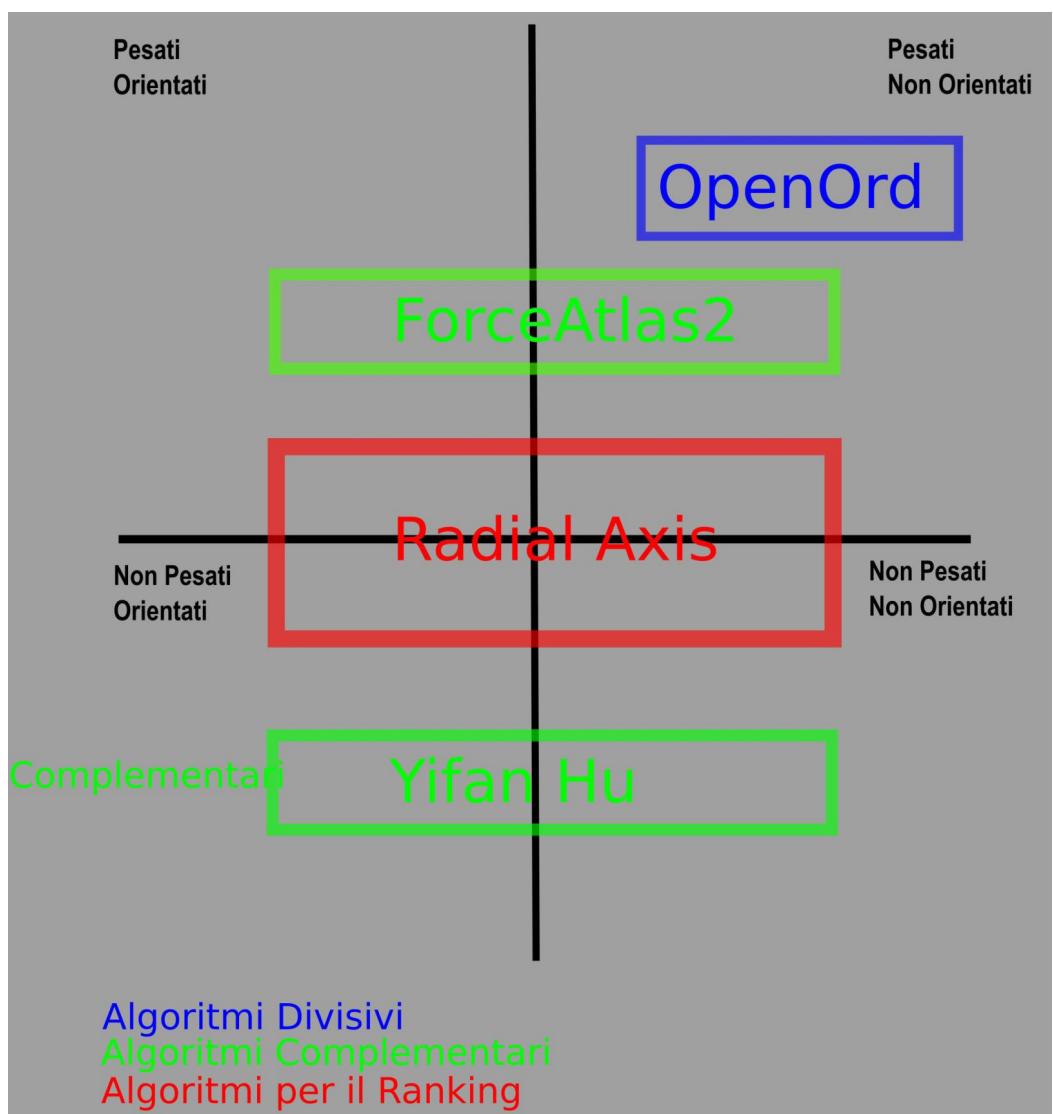



Fig. 03. Algoritmi di layout in funzione delle loro capacità operative. Elaborazione Amedeo Ganciu

Tab. 4. Tabella riassuntiva con analisi dei Nodi, scomposizione delle singole reti in Community e qualità dei cluster.

Rete	Nodi	Archi	Community	Modularity
Autori	104'207	189'918	30'073	0.883
Keywords	95'942	569'779	1'859	0.612
Citazioni tra articoli	1'361'821	1'853'915	41'757	0.838

La mappatura delle reti attraverso differenti algoritmi può agevolare la comprensione delle dinamiche e relazioni interne in ciascuna rete.
Questi strumenti sono stati scelti considerando la loro capacità analitica e di rappresentazione in funzione delle diverse caratteristiche delle reti: pesate/non pesate, orientate/non orientate, numero di nodi, tempi di risposta più brevi possibili (tab.5).

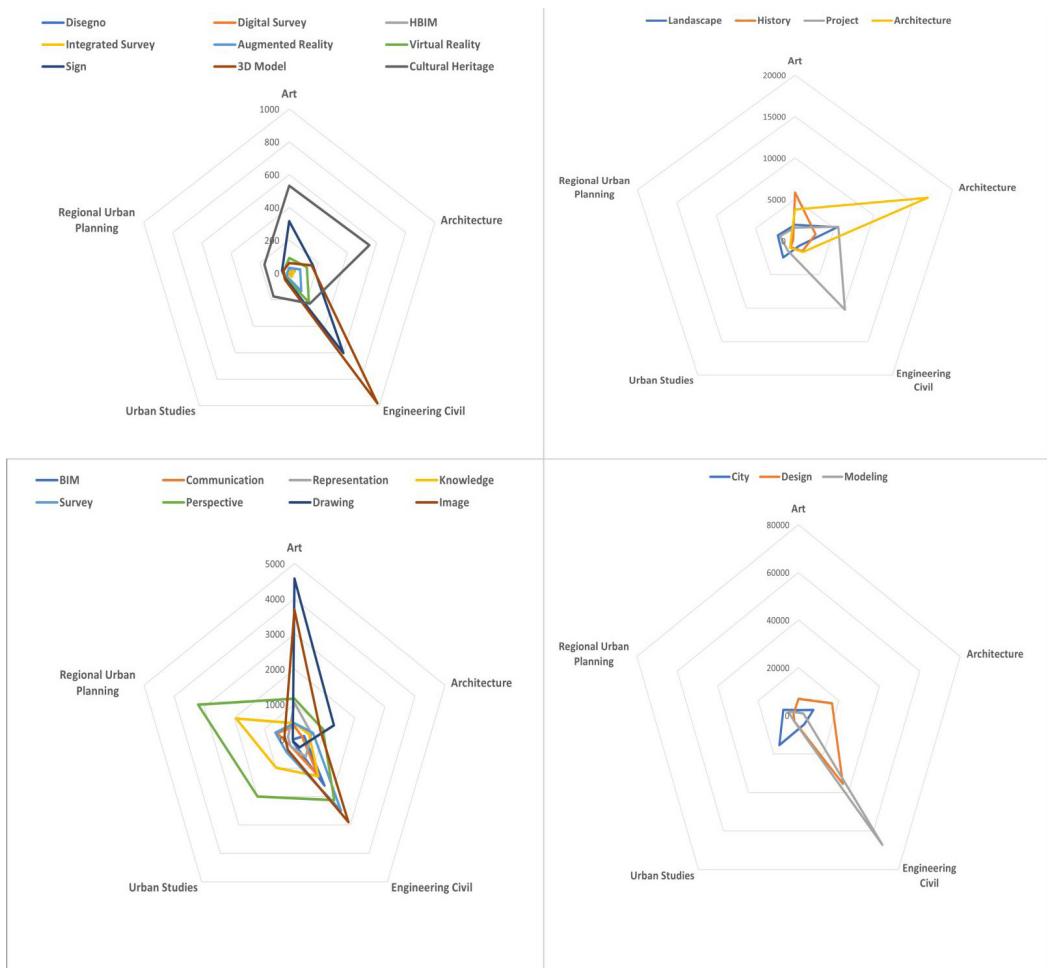



Fig. 04. Distribuzione degli articoli scaricati per categorie disciplinari di WoS. Elaborazione Amedeo Ganciu.

La visualizzazione della rete degli autori conferma i risultati numerici per complessità elevata e frammentazione delle collaborazioni tra i ricercatori, osservabile con sufficiente efficacia attraverso l'applicazione del layout divisivo OpenOrd ed evidenziando con colori differenti ciascuna delle oltre 30'000 famiglie di autori, ossia i ricercatori che tendono a collaborare tra loro più frequentemente (fig. 05a).
L'algoritmo complementare *ForceAtlas2* non offre in questi casi risultati grafici ugualmente soddisfacenti per la comprensione della elevata frammentazione (fig. 05b).

Tab. 5. Scelta dei differenti algoritmi di layout per le reti.

	Reti		
Algoritmo	Autori	Keywords	Citazioni tra articoli
OpenOrd	Si	Si	No
Force Atlas 2	Si	Si	No
Yifan Hu	No	No	Si
Radial Axis	Si	Si	Si

Tuttavia, lavorando con una rete di autori più pulita, cioè filtrandola considerando quelli con il maggiore numero di collaborazioni ed ottenendo una rete di "solo" 25'974 nodi e 80'006 connessioni, il layout *ForceAtlas2* sviluppa una elaborazione grafica più interpretabile, consentendo di osservare il comportamento delle diverse famiglie di Autori (fig. 06).
Un ulteriore zoom sempre attraverso lo stesso layout rivela le collaborazioni tra i 130 autori principali all'interno delle dieci community principali, cioè quelle composte dal maggiore numero di autori (fig. 07b).
L'algoritmo di *ranking Radial Axis* consente ovviamente di discretizzare l'importanza relativa di ciascun autore all'interno di ogni gruppo di ricerca, ma limita la comprensione delle collaborazioni tra i diversi gruppi di ricerca (fig. 07a).
Il particolare risultato ottenuto potrebbe anche essere dovuto alla evidente tendenza della rete verso i settori ingegneristici, in particolare quelli connessi allo sviluppo di nuovi strumenti tecnologici per la rappresentazione, settore in cui da diversi anni l'area asiatica sta marcatamente segnando la sua futura evoluzione.
L'applicazione degli algoritmi di visualizzazione alla rete delle keyword consente di comprendere quali sono i topics di ricerca più caldi e come sono connessi tra loro, ovviamente i diversi algoritmi hanno fornito risultati differenti sul profilo della comprensione grafica. In questo caso, la complessità della rete è pure superiore per numerosità delle connessioni tra i nodi, infatti sia *l'OpenOrd* (fig. 08a) che il *ForceAtlas2* (fig. 08b) non hanno consentito elaborazioni apprezzabili, cioè che consentano una efficace interpretazione della struttura. Come nel caso precedente un filtro sulle keyword più frequenti permette di ridurre la complessità della rete per apprezzare meglio le differenze tra i due strumenti di layout con l'algoritmo *OpenOrd* (fig. 09b), che consente di enfatizzare meglio le differenze tra la composizio-

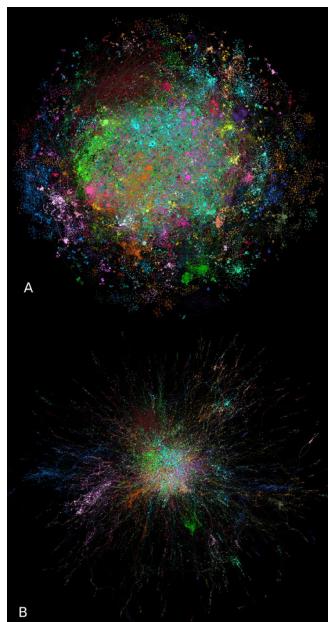
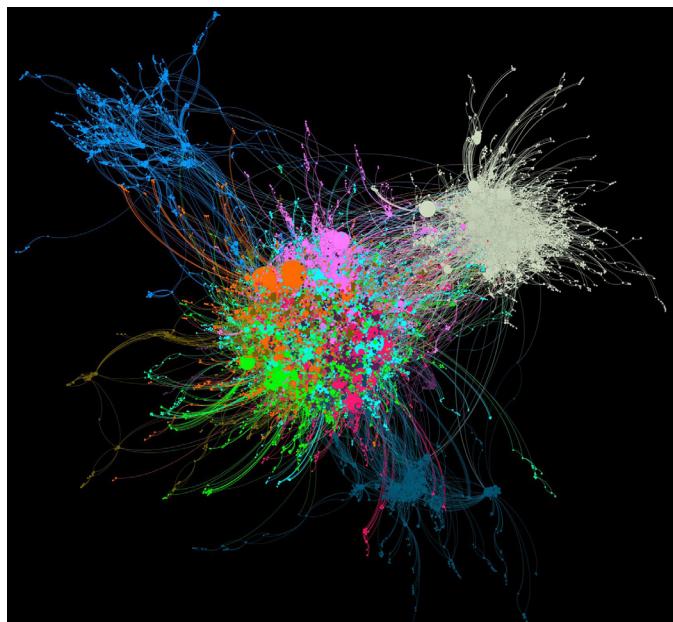



Fig. 05. La rete di collaborazione tra Autori visualizzata con OpenOrd (A) e con ForceAtlas2 (B), ogni nodo rappresenta un autore e la sua dimensione e proporzionale alla sua attività di ricerca, gli archi tra i nodi indicano la collaborazione tra due autori, i nodi dello stesso colore appartengono alla stessa community. Elaborazione Amedeo Ganciu.

Fig. 06. La rete di collaborazione tra Autori filtrata e visualizzata con ForceAtlas2, ogni nodo rappresenta un autore e la sua dimensione e proporzionale alla sua attività di ricerca, gli archi tra i nodi indicano la collaborazione tra due autori, i nodi dello stesso colore appartengono alla stessa community. Elaborazione Amedeo Ganciu.

ne delle community e le relazioni che si instaurano tra queste rispetto al *ForceAtlas2* (fig. 09a). Con l'ulteriore zoom sulle keywords principali contenute nelle community più popolose è possibile individuare alcune tematiche di ricerca che caratterizzano la rete.
Si può osservare come gli argomenti del "Design" sono fortemente connessi allo studio delle strutture sotto diversi aspetti e attraverso "Simulation" si congiunge all'altra community di colore magenta dominata apparentemente dal "BIM" e dal "Project Management" (fig. 10). Nella terza rete, quella delle citazioni, l'elevato numero di community consente una efficace visualizzazione della frammentazione anche considerandola nella sua interezza, ossia visualizzando oltre 1'000'000 di nodi (fig. 11). Ad un successivo approfondimento con i soli algoritmi di layout *ForceAtlas2* (fig. 12A) e Yifan Hu (fig. 12B) è possibile osservare meglio la struttura e connessione tra le famiglie di articoli, cioè gruppi di pubblicazioni che tendono a citarsi tra loro molto frequentemente.

Conclusioni

La ricerca presentata è stata finalizzata a tradurre in immagini dati e informazioni difficilmente analizzabili e interpretabili. In particolare, l'obiettivo è stato elaborare visualizzazioni utili alla comprensione della configurazione e delle relazioni che definiscono la comunità scien-

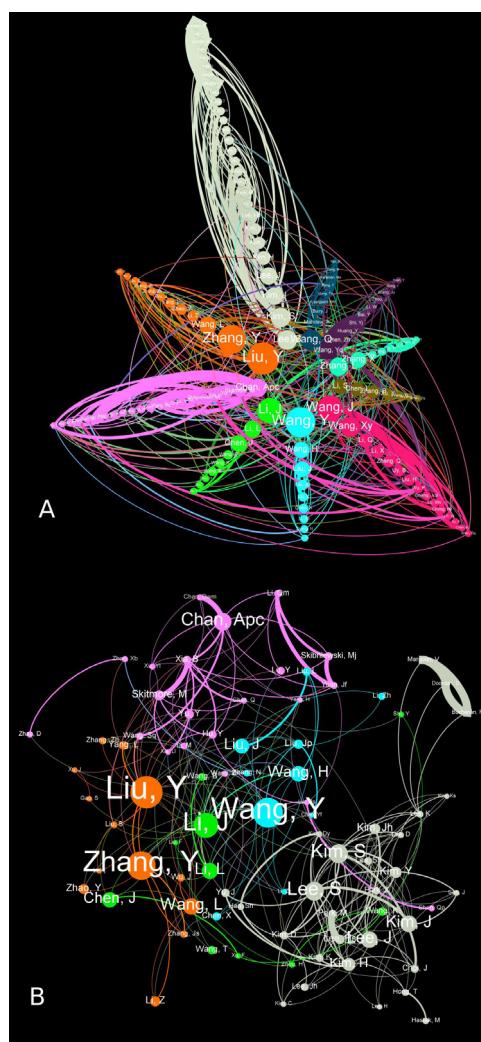



Fig. 07. La rete di collaborazione tra i principali Autori delle Community più importanti visualizzata con RadialAxis(A) e con ForceAtlas2, ogni nodo rappresenta un autore e la sua dimensione e proporzionale alla sua attività di ricerca, gli archi tra i nodi indicano la collaborazione tra due autori, i nodi dello stesso colore appartengono alla stessa community. Elaborazione Amedeo Ganciu.

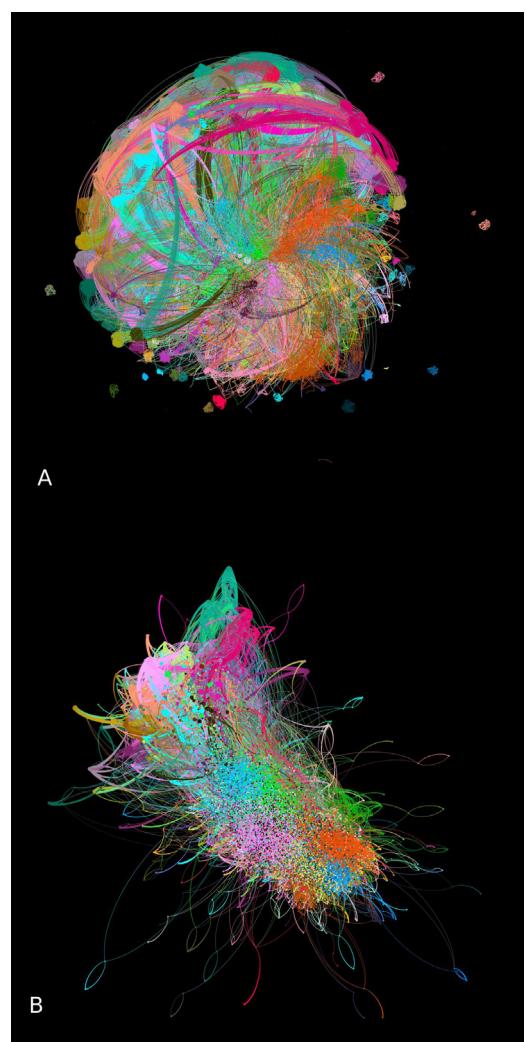
Fig. 08. Rete di citazioni tra keywords delle pubblicazioni scaricate visualizzate con OpenOrd (A) e ForceAtlas2 (B), ogni nodo rappresenta un keywords di un articolo e la sua dimensione e proporzionale alla sua frequenza nella rete, gli archi tra i nodi indicano la copresenza delle keywords sulla stessa pubblicazione, i nodi dello stesso colore appartengono alla stessa community. Elaborazione Amedeo Ganciu.

tifica internazionale che si addensa attorno alle scienze grafiche a partire dalle keywords, che definiscono i principali ambiti di ricerca. I risultati permettono di evidenziare alcune peculiarità utili a interpretare le relazioni tra i diversi ricercatori, i diversi temi di indagine e i diversi prodotti della comunità scientifica. Dalla distribuzione degli articoli scaricati per categorie disciplinari di *WoS* (fig. 05) si può evidenziare le associazioni tra parole chiave e macro ambiti del sapere. Queste associazioni sono utili a collocare anche gli interessi di ricerca,

Figura 09. Rete delle keywords filtrata e visualizzata con ForceAtlas2 (A) e con OpenOrd (B), ogni nodo rappresenta un keywords di un articolo e la sua dimensione e proporzionale alla sua frequenza nella rete, gli archi tra i nodi indicano la copresenza delle keywords sulla stessa pubblicazione, i nodi dello stesso colore appartengono alla stessa community. Elaborazione Amedeo Ganciu.

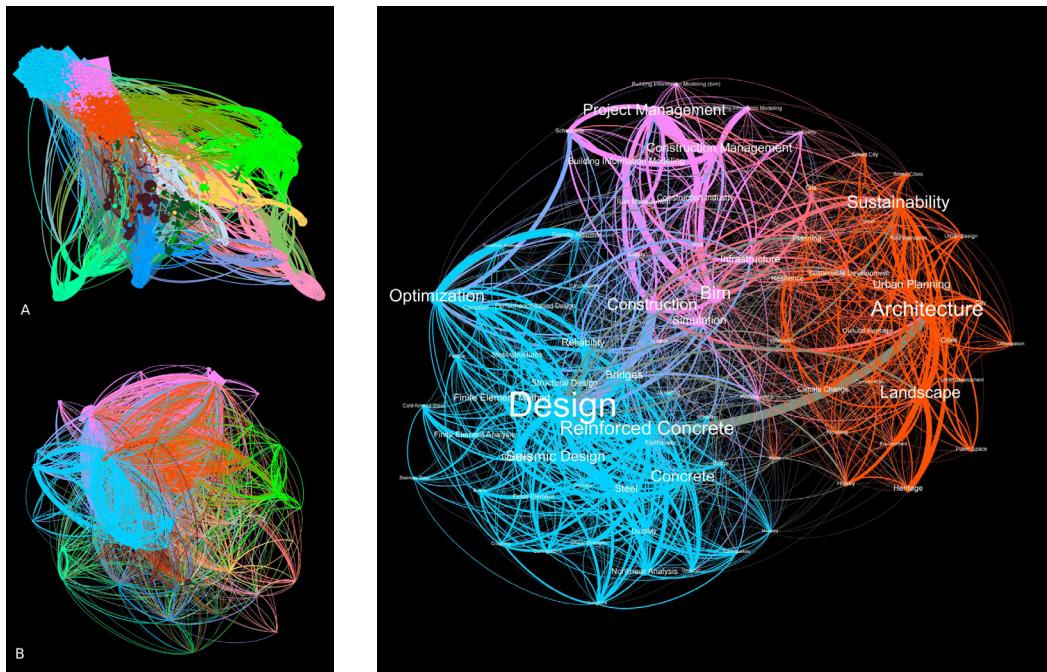
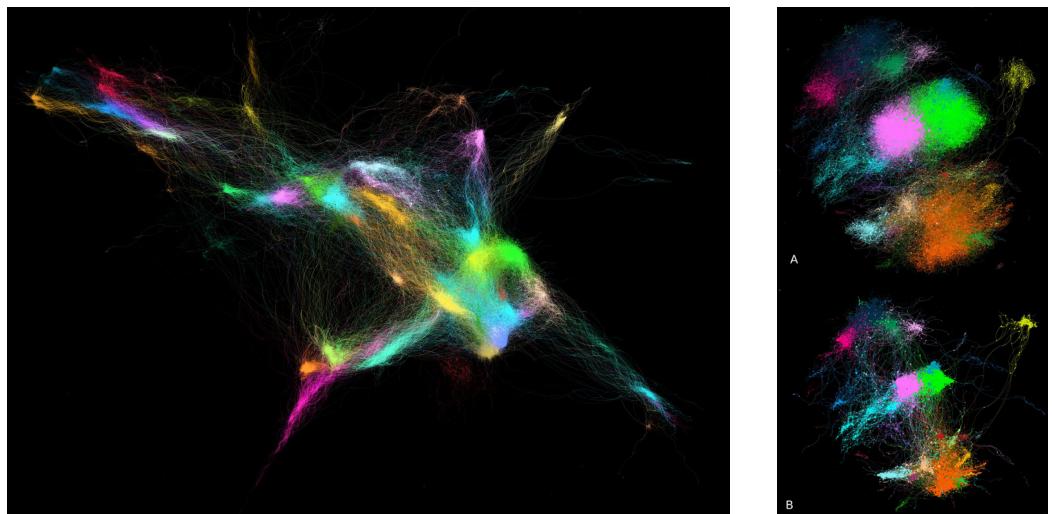
Fig. 10. Particolare della rete delle keywords visualizzata con OpenOrd, ogni nodo rappresenta un keywords di un articolo e la sua dimensione e proporzionale alla sua frequenza nella rete, gli archi tra i nodi indicano la copresenza delle keywords sulla stessa pubblicazione, i nodi dello stesso colore appartengono alla stessa community. Elaborazione Amedeo Ganciu.

sintetizzati dalle keywords, all'interno degli ERC. Ad esempio, è possibile notare come, nel dibattito internazionale, le ricerche focalizzate sulla modellazione (3D modeling) ricadano all'interno della sfera di interesse strettamente ingegneristica, mentre la ricerca sul disegno (drawing) sia maggiormente associata alla sfera delle arti. Nella visualizzazione della rete di collaborazione tra i principali Autori delle Community più importanti (fig. 08) emergono degli addensamenti di relazioni interpretabili come comunità citazionali che meriterebbero ulteriori approfondimenti utili alla comprensione del sistema citazionale bibliometrico. La visualizzazione della rete delle keywords (fig. 11) consente invece di visualizzare gli addensamenti tematici all'interno dei prodotti scientifici e definire una nebulosa di interessi di ricerca della comunità scientifica che opera nel campo dei settori disciplinari definiti da *WoS*, all'interno dei quali può riconoscersi la comunità scientifica del disegno. Infine, l'applicazione della Social Network Analysis per lo studio del panorama scientifico e la successiva visualizzazione attraverso diversi algoritmi di layout si dimostra un valido ed efficace strumento di indagine. I risultati delle analisi ottenute dal massivo download di pubblicazioni da *WoS* sono approssimativamente in linea con ciò che si può ipotizzare di associare alla disciplina delle scienze grafiche, tuttavia è bene ricordare nuovamente che per il caso italiano la maggior parte della produzione scientifica non essendo catalogata attraverso strutture bibliometriche è sfuggita a questa ricerca, quindi al fine di comprendere come il "settore" italiano collabora alla produzione internazionale, nelle prossime ricerche sarà necessario operare precedentemente un minuzioso lavoro di raccolta delle informazioni per poi trasformarle nel formato ISI.file e quindi inglobarle all'interno delle strutture illustrate in questa ricerca.

Fig. 11. Rete di citazione tra articoli scientifici visualizzata con algoritmo di layout ForceAtlas2, ogni nodo rappresenta una pubblicazione, gli archi tra nodi indicano una relazione di citazione, considerata la complessità non è possibile osservare a questa scala gerarchie tra articoli, i nodi dello stesso colore appartengono alla stessa community. Elaborazione Amedeo Ganciu.

Fig. 12. Rete di citazione tra articoli visualizzata con ForceAtlas2 (A) e Yifan Hu (B), ogni nodo rappresenta una pubblicazione, la sua dimensione è proporzionale al numero di citazioni ricevute, gli archi tra nodi indicano una relazione di citazione, i nodi dello stesso colore appartengono alla stessa community. Elaborazione Amedeo Ganciu.

Riferimenti bibliografici

Abt, H. A. (1998). Why some papers have long citation lifetimes. In *Nature*, n. 395(6704), pp.756-757.

Balestrieri, M., & Ganciu, A. (2019, July). The Use of Graphs to Explore the Network Paradigm in Urban and Territorial Studies. In *International and Interdisciplinary Conference on Image and Imagination*, pp. 1120-1132. Springer, Cham.

Batagelj, V. (2003). Efficient algorithms for citation network analysis. *arXiv preprint* cs/0309023.

Blondel, V. D., (2008). Fast unfolding of communities in large networks. In *Journal of statistical mechanics: theory and experiment*, 2008(10), P10008.

Boccaletti, S., (2006) Complex networks: Structure and dynamics. *Physics Reports*, No. 424, pp. 175–308.

Bocconcino, M., Vozzola, M. Infografica e visualizzazione grafica: nuovi modi per la rappresentazione dei dati. In *DIENNE V.7*, 7/2020.

Börner, K., Chen, C., Boyack, K. W. (2003). Visualizing knowledge domains. In *Annual review of information science and technology*, 37(1), 179-255.

Caldarelli, G. (2007). Scale-free networks: complex webs in nature and technology. In *Oxford University Press*.

Cicalò Enrico (2020a). Connessioni tra saperi: disciplinarietà, interdisciplinarità e transdisciplinarietà delle scienze grafiche/ Connections between knowledge: disciplinarity, interdisciplinarity and transdisciplinarity of graphic sciences. In Arena A., "et al." (a cura di). *Connettere. Un disegno per annodare e tessere.* Atti *del 42° Convegno In- ternazionale dei Docenti delle Discipline della Rappresentazione/Connecting. Drawing for weaving relationships. Proceedings of the 42th International Conference of Representation Disciplines Teachers*. Milano: FrancoAngeli, pp. 316-337.

Cicalò, E. (2020b). Exploring Graphic Sciences. In *International and Interdisciplinary Conference on Image and Imagination* (pp. 3-14). Springer, Cham.

Cicalò, E., Menchetelli, V., (2020) Informazioni accessibili. la visualizzazione dati nell'epoca dei big data, degli open data e degli open tools. In *DIENNE V.7*, 7/2020.

Cicalò, E., Pileri, M., Valentino, M., (2021). Connessione tra saperi. Il contributo delle scienze grafiche nella ricerca in ambito medico. Connecting Knowledge. The Contribution of Graphic Sciences to Medical Research. In A. Arena, M. Arena, D. Mediati, P. Raffa, *42° Convegno Internazionale dei Docenti delle Discipline della Rappresentazione CONNETTERE_un disegno per annodare e tessere Linguaggi, Distanze, Tecnologie*. 16, 17 e 18 settembre 2021, pp. 510-527. Reggio Calabria e Messina.

Cook, E.A. (2002) Landscape structure indices for assessing urban ecological networks. In *Landscape. Urban Planning.* 58, pp.269–280.

De Meo, (2011, November). Generalized louvain method for community detection in large networks. *In 11th international conference on intelligent systems design and applications*, pp. 88-93. IEEE.

Derrible, S.; Keendy, C. (2011) Applications of Graph Theory and Network Science to Transit Network Design. In *J. Transp. Rev.*, No. 31, pp. 495–519.

Estrada, E., Bodin, Ö. (2008) Using network centrality measures to manage landscape connectivity. *In Ecol. Appl.*, No. 18, pp. 1810–1825.

Ganciu, A.; Balestrieri, M.; Cicalò, E. (2016, June) Visualising the research on visual landscapes. Graph representation and network analysis of international bibliography on landscape. In *Proceedings of the XIV International Forum Le Vie dei Mercanti*, Capri, Italy, 16–18 June 2016.

Ganciu, A. (2018). Dynamics of metropolitan landscapes and daily mobility flows in the Italian context. An analysis based on the theory of graphs. In *Sustainability*, Vol. 10(3), p. 596.

Garfield, E., Sher, I. H., Torpie, R. J. (1964). The use of citation data in writing the history of science. Institute for Scientific Information Inc Philadelphia PA.

Gigliarelli, E. (2020) La rappresentazione della conoscenza utilizzando le mappe cognitive per favorire la ricerca collaborative. In *Building Information Modeling, Data & Semantics*, 7/2020.

Green, N. (2007). Functional polycentricity: A formal definition in terms of social network analysis. *In Urban Stud*., No. 44, pp.2077–2103.

Hu, Y.F., (2005) Efficient and high quality force-directed graph drawing. In *The Mathematica Journal*, 10 (37-71), 2005.

Jin, J. H., Park, S. C., Pyon, C. U. (2011). Finding research trend of convergence technology based on Korean R&D network. In *Expert Systems with Applications*, 38(12), 15159-15171.

Lo, H.K.; Tung, Y.K. (2003) Network with degradable links: Capacity analysis and design. In *Transp. Res. Part B Methodol.* 37, pp. 345–363.

Luigini, A. (2021). Ricerca interdisciplinare e ICAR17: una proposta per la definizione di un modello condiviso. *42° Convegno Internazionale dei Docenti delle Discipline della Rappresentazione CONNETTERE_un disegno per annodare e tessere Linguaggi, Distanze, Tecnologie*. 16, 17 e 18 settembre 2021, Reggio Calabria e Messina.

Luigini A (2021b); Riviste scientifiche nel settore ICAR17: analisi quantitativa delle keywords e dei temi di ricercar. In A. Arena, M. Arena, D. Mediati, P. Raffa, *42° Convegno Internazionale dei Docenti delle Discipline della Rappresentazione CONNETTERE_un disegno per annodare e tessere Linguaggi, Distanze, Tecnologie*. 16, 17 e 18 settembre 2021, pp. 510-527. Reggio Calabria e Messina.

Martin S. (2011) OpenOrd: An Open-Source Toolbox for Large Graph Layout. In *SPIE Conference on Visualization and Data Analysis (VDA)*.

Mathis, P. (2007). Graphs and Networks: Multilevel Modeling.

Newman, M. E. (2003). The structure and function of complex networks. SIAM review, 45(2), pp. 167-256.

Newman, M.E.J.; Girvan, M. (2004) Finding and evaluating community structure in networks. In *Phys. Rev. E.* 69, pp. 1–15.

Patuelli, R. (2007) Network analysis of commuting flows: A comparative static approach to German data. In *Netw. Spat. Econ.* No.7, pp. 315–331.

Porta, S.; Crucitti, P.; Latora, V. (2006). The network analysis of urban streets: A dual approach. Phys. A Stat. Mech. Its Appl. No. 369, pp. 853–866.

Reuters, T. (2016). Web of knowledge-real facts-ip & science-thomson reuters.

Robins, G. (2013). A tutorial on methods for the modelling and analysis of social network data. In J. Math. *Psychol*. No. 57, pp 261–274.

Salas-Olmedo, M.H.; Nogués, S. (2012). Analysis of commuting needs using graph theory and census data: A comparison between two medium-sized cities in the UK. In *Appl. Geogr*. No. 35, pp. 132–141.

Urban, D.L.; Minor, E.S.; Treml, E.A.; Schick, R.S. (2009) Graph models of habitat mosaics. In Ecol. Lett. No. 12, pp.260–273.

Xie, F.; Levinson, D. Evolving (2011) Transportation Networks. Berlin:Springer Science & Business Media.

Autori

Amedeo Ganciu, Dipartimento di Architettura, Design e Urbanistica, Università degli Studi di Sassari, aganciu@uniss.it
Andrea Sias, Dipartimento di Architettura, Design e Urbanistica, Università degli Studi di Sassari, a.sias19@studenti.uniss.it

Per citare questo capitolo: Ganciu Amedeo, Sias Andrea (2022). Visualizzare la conoscenza. La rappresentazione delle reti citazionali internazionali nell'ambito delle scienze grafiche/Visualising the knowledge. The representation of international citation networks in the graphic sciences. In Battini C., Bistagnino E. (a cura di). *Dialoghi. Visioni e visualità. Testimoniare Comunicare Sperimentare. Atti del 43° Convegno Internazionale dei Docenti delle Discipline della Rappresentazione/Dialogues. Visions and visuality. Witnessing Communicating Experimenting. Proceedings of the 43rd International Conference of Representation Disciplines Teachers*. Milano: FrancoAngeli, pp. 2479-2502.

Isbn 9788835141938

43rd International Conference
of Representation Disciplines Teachers
Congress of Unione Italiana Per Il Disegno - Proceedings 2022

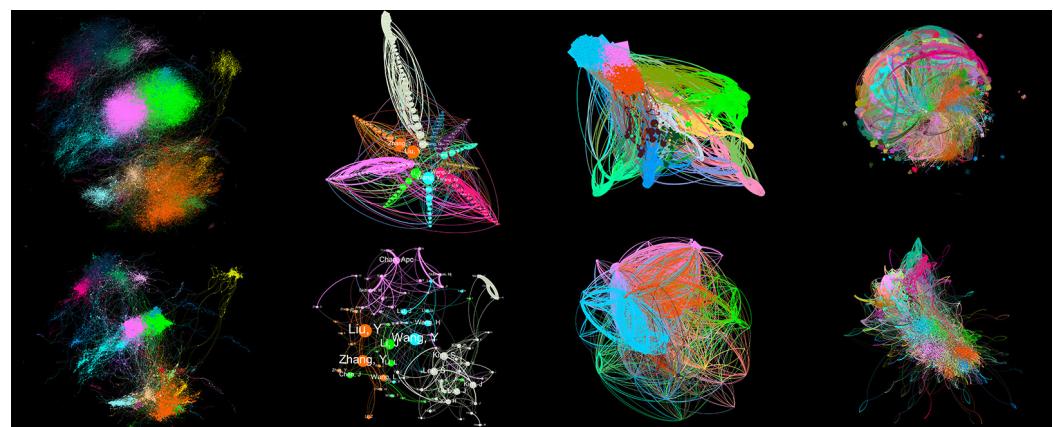
Visualising the knowledge. The representation of international citation networks in the graphic sciences

Amedeo Ganciu
Andrea Sias

Abstract

This paper explores the visualization of big data related to citation networks inside the research fields of interest of graphic sciences. Starting from previous researches carried out in the last years also during UID conferences by professors afferent to the ICAR/17 sector, through a massive search within the Web of Science database, it was possible to identify and analyze numerically and graphically more than 100'000 documents in order to reconstruct the citation networks among the articles, the collaborations between Authors, and the citation networks between the keyswords present on each publication. In this paper we test the application of Social Network Analysis tools to investigate and map the interaction between fields of knowledge and the researchers who represent them. Aware of some limitations inherent in the nature of research, this study is configured as a first experiment that applies methods and tools established in the international literature to a new case study. The research can not be considered complete because it refers only to bibliometric data from a single database (*WoS*) that would need to be integrated with data from other databases (*Scopus*) and data on non-bibliometric research products that currently are prevalent within the scientific community ICAR/17.

Keywords
Data visualization, big data, graphs, citation networks, graphic sciences

Topic
Mapping

Graphic processing of the networks. (elaboration Amedeo Ganciu).

doi.org/10.3280/oa-832-c155

Introduction

Inter- and trans-disciplinary research is gaining a lot of interest among researchers for many reasons and yielding different results [Cicalò 2020a], not only because it makes possible a unified vision in which disciplinary boundaries are taken as a reference and not as a limit [Luigini 2021a], but also for its economic potential associated with a contamination and transfer of knowledge between disciplines [Jin et al. 2011]. Among the techniques for representing the complexity of big data, several researchers are highlighting the centrality of infographics and data visualisation [Bocconcino and Vozzola 2020, Cicalò and Menchetelli 2020; Gigliarelli et al. 2020]. In particular, among the instruments supporting the knowledge of complex dynamical systems, the modelling addressed to the reticular paradigm is acquiring increasing attention from the international scientific community with numerous application examples [Newman and Girvan 2004; Boccaletti et al. 2006], in economic and social systems [Robins 2013; Ganciu et al 2016; Green 2007], transport and mobility infrastructures [Ganciu et al. 2018; Porta et al. 2006; Xie and Levinson 2011; Derrible and Keendy, 2011], urban and spatial planning [Cook 2002; Estrada and Bodin 2008; Urban 2009], providing the tools for a better understanding of the phenomenon [Balestrieri and Ganciu 2019]. Newman (2003) identifies four classes of networks: social networks, such as forms of contact or interaction between individuals or academic citation networks; information networks, such as links in the World Wide Web; technological networks, such as water, transport and energy systems; and biological networks, such as food networks with prey and predators. In all four cases, the methodological background for the analysis and representation of networks can be attributed to the development of graph theory by Euler in 1736, who was the first to lay the fundamentals of the discipline known as graph theory in solving the Konigsberg Bridge problem (fig. 01). In the social network category, bibliographic networks are an example of a non-spatialised network, in which nodes' and arcs' Euclidean position loses its significance to the disadvantage of the importance assumed by the nodes through the relationships they establish with each other, i.e. the number of citations they receive. The amount of citations a paper receives from other papers is the factor that determines its permanence or transience in the visibility window of a citation network, as was already demonstrated in pioneering work in the 1970s [Garfield et al., 1964; Batagelj, 2003; Börner et al, 2003; Luigini, 2021b], and yet, within scientific cyberspace it is still possible to identify recurrent citation structures that are based on papers with a high number of citations, encompassed by other papers, which tend to be marginal and transient in nature in the scientific landscape [Abt 1998]. In this particular study, previous works have already demonstrated how the field of Design is characterised by a high level of heterogeneity at a

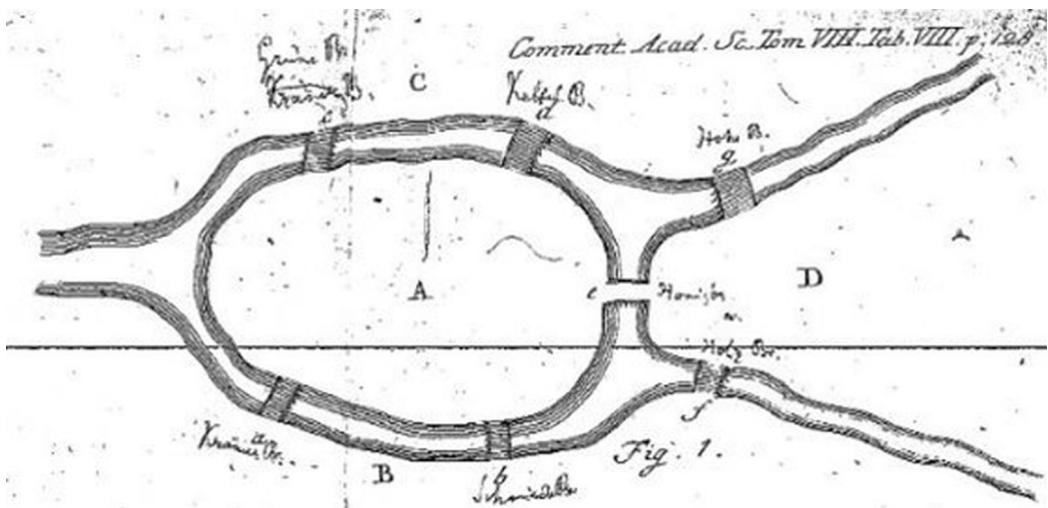



Fig. 01. Map of the town of Konigsberg made by Euler to solve the problem.

National level [Luigini 2021a and 2021b], in which contributions from the hard and humanities sciences coexist. Using these papers as a starting point, we proceeded to carry out a research within the Web of Science database of articles and documents attributable to the graphic sciences [Cicalò 2020b]; in particular, it was possible to identify and analyse numerically and graphically over 100,000 documents that allowed us to reconstruct the citation networks between articles, the collaborations between authors, and the citation networks between the keywords present in each publication. By analysing the citations between articles, the distribution of publications is reconstructed, identifying the most significant works, i.e. those with the highest number of citations, and which are potentially most likely to influence the direction of subsequent research. By the network of authors it is possible to measure and visualise how scholars collaborate with each other and at what intensity; lastly, with the network of citations of research keywords it is possible to see which are the main themes and how they are connected. The research also tests the application of different algorithms for the visualisation and automated layout of the networks, which is absolutely necessary considering the complexity of the networks in terms of nodes and arcs, obtaining different results in terms of simplicity of reading and interpretation of the results.
This study does not pretend to be exhaustive of the whole discipline due to various limitations, primarily the lack of bibliometric relevance of the Italian sector; although several journals in the sector are indexed by systems such as Web of Science or Scopus [Luigini, 2021b], a significant number of papers were published in journals that were not censored on the bibliometric profile and therefore escape this analysis. For this reason, the study should be considered as a first element towards a future, more integral research, more focused on the national panorama, which aims at highlighting the Italian role within the international debate.

Method and Data

Data for the analysis were taken from the *Web of Science* (*WoS*) online database currently maintained by *Clarivate Analytics*, which indexes papers published in over 12,000 journals and 160,000 conference proceedings, totalling over 90 million articles and over a billion cited references, [Reuters 2016]. The selection of the articles was carried out by looking for the presence in their title of words considered significant within the disciplinary sector and as already mentioned used in previous research such as by Luigini for the 42nd International Conference of Teachers of the Disciplines of Representation Congress of the Italian Union for Drawing (tab.1). The word "Drawing", although not explicitly mentioned in the previous research, was deliberately added in view of its use in English-language literature. Then, for each voice, a further selection and refinement was made using the disciplinary categories indicated within the *WoS* that are most relevant to the disciplinary field of graphic sciences (tab. 2). Thus, through the cross-query: Select = ["Word in the Title" AND "WoS Category"] a total of 260'002 records were initially found.

Word in the title	N°	Word in the title	N°	Word in the title	N°
Modeling	72753	Survey	3813	Integrated Survey	63
Design	59567	Knowledge	3112	HBIM	58
City	28266	Representation	2362	Digital Survey	48
Architecture	22218	Communication	2053	Disegno	32
Project	18242	BIM	1920		
History	11211	Cultural Heritage	1291		
Landscape	10219	3D Model	1191		
Image	7630	Sign	1092		
Drawing	6154	Virtual Reality	449		
Perspective	6010	Augmented Reality	248	Totale	260002

Tab. 1. Words used to search for scientific publications within WoS.

Tab. 2. WoS subject categories used to filter publications.

Selected subject categories within Web of Science	
Art	Urban Studies
Architecture	Regional Urban Planning
Engineering Civil	

In order to avoid analysing duplicate records as two, three or more words may appear simultaneously within the same title, leading to double, triple etc. counting of information, the analysis base was refined by uniquely identifying each publication and eliminating duplicate records. This extremely delicate and fundamental procedure aimed at guaranteeing the correctness of any subsequent analysis was carried out through the automated reading of the DOI (Digital Object Identifiers) code unique to each record present in the *WoS*, allowing us to obtain a 'clean' database of 101'741 records, i.e. the publications used in this research. For each of the selected records, the corresponding ISI.file was downloaded from the *WoS*, which is a kind of metadata sheet, containing all the main information of the paper such as the authors, the abstract, the keywords, the journal or book in which it was published, and a crucial aspect for citation networks the documents it cites among its bibliography entries; each of this information is classified within the ISI.file through TAGs that bibliometric systems recognise and can use. The overall database, thus consisting of 101,741 ISI.files, was pre-processed within the Sci-2Tool platform, created by the Cyberinfrastructure for the Network Science Center - Indiana University (USA). In particular, the networks of "Authors", "Keywords" and "Citations between publications" were modelled, differing in terms of search purpose, complexity in terms of number of nodes and arcs, and type in terms of orientation and weight of the arcs (tab.3). In particular, the networks of "Authors", "Keywords" and "Citations between publications" were modelled, differing in terms of search purpose, complexity in terms of number of nodes and arcs, and type in terms of orientation and weight of the arcs (tab.3).

Tab. 3. Networks analysed with indication of type.

Network	Network orientation	Weight of the edges
Authors	Not oriented	Yes
Keywords	Not oriented	Yes
Citations between articles	Oriented	No

A network can be represented through its mathematical formalisation; through its graphical representation; or through its matrix form (fig. 02).
In terms of the mathematical discipline and its applications, existing algorithms make it possible to calculate the mutual influence between different nodes in the network by computing arcs with different weights and directions [Mathis 2007].
For example, a graph is defined as oriented if there are obligations or rules that determine the direction or walkability of an arc (as in a one-way street), a graph can be weighted if it is possible to assign a different importance to each arc, as an expression of the intensity of the relationships between nodes [Caldarelli 2007].
In the Authors' network, each node represents an author, and the existence of an arc between two nodes indicates their collaboration on a publication; as the frequency with which this event occurs increases, so does the weight of the arcs; hence, the force that somehow binds them together; the network is not oriented, since saying that "A" collaborates with "B" is the same as saying that "B" collaborates with "A".

Fig.02. Example of the three representations of the same graph. Processing by Amedeo Ganciu.

Using the 'Keywords' network, it is possible to extrapolate and represent information on which research topics are most attractive to researchers and how they are related to each other through the citation relationships between articles; this network is also not oriented because, as in the previous case, it would not have meant directionality, while it did mean giving a weight to the arcs as an indicator of the frequency with which two keywords are connected by a citation relationship.
The network of "Citations between articles" obviously indicates the relations between the publications, i.e. "who cites whom", all the arcs have equal importance, so the network is not weighted, but, contrary to the previous cases, it is oriented, because even intuitively it is logical to think that if article "A" cites article "B", it is not possible that the opposite happens at the same time.
After the three networks had been constructed, basic quantitative analyses of the structures were carried out using the open source software Gephi. In order to obtain the earliest information on their morphology, the node degree (Deg) was calculated for each network, i.e. the sum of the number of arcs converging on each node of the network [Foulds 2012], and in the case of weighted networks, the sum of all the weights of all the arcs converging on each node of the network. In addition, in all three cases, analysis was carried out to identify communities through the application of Louviain's algorithm [Blondel et al. 2008].
The last analysis makes it possible to determine the presence of clusters within a network, which are characterised by having the highest density of relations (arcs) within them and at the same time the lowest density between nodes belonging to different communities [De Meo et al. 2011].
Lastly, different algorithms were applied to visualise the networks belonging to the typology that emphasise the divisions, complementarities and ranking between the nodes that make up the networks. In particular, the 'OpenOrd' algorithm [Martin et al. 2011] was used to visually maximise the sub-structures or communities that make up a network.
This algorithm expects an undirected, weighted network and can process network representations with between 100 and 1,000,000 nodes. Among the complementarity visualisation algorithms, we chose the "*ForceAtlas2*", native within Gephi and developed by Mathieu Jacomy, and the Yifan Hu Multilevel layout developed in 2005 [Hu. 2005].
However, ForceAtlas2 allows the user to choose whether to use the weight of the arcs to obtain a better visualisation of the network, which Yifan Hu Multilevel layout does not. Finally, among the algorithms that emphasise the ranking we chose to use the Radial Axis Layout, native within Gephi and developed by Matt Groeninger, allows us to create orders for elements characterised by homophily properties such as belonging to the same community or other characteristics, it is also capable of managing networks with a complexity of up to 1,000,000 nodes (fig. 03).

Results

The first phase of screening carried out within the *WoS* portal indicates some of the words searched within the titles of the publications as significantly preponderant in terms of statistical numerosity, for example "Modeling" and "Design" present in several tens of thousands of publications, while other terms such as "Design" and "Digital Survey" significantly more marginal compared to the former. If we consider the disciplinary categories used in *WoS* to index publications, we can see that the civil engineering and architecture class is predominant for the majority of searches, particularly those containing the terms 'Modeling', 'Design', 'Project' and 'Architecture' in the title. The classes of territorial and urban planning include instead most of the documents containing the words "Perspective" and "Knowledge". Finally, in most cases titles containing the words "Drawing" and "Image" are classified by *WoS* as belonging to the class Art. (fig. 04).
The extraction of the three networks from the ISI.file has generated structures that are significantly different in complexity in terms of the number of nodes and connections, but quite similar in their distributional morphology, made up of a different number of communities that however present a quality in their separation expressed through the Modularity value that is very high for all three cases (tab.4).

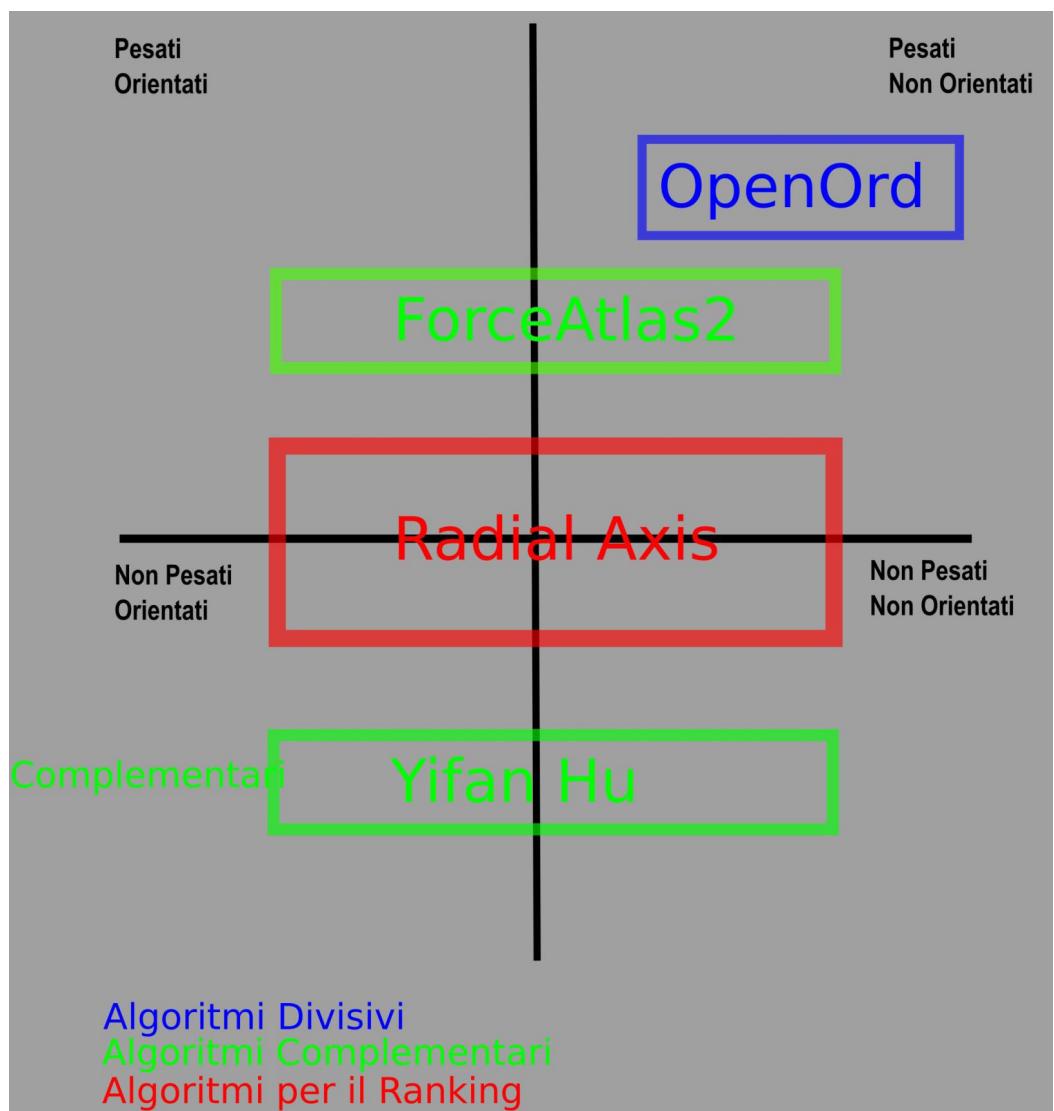



Fig. 03. Layout algorithms as a function of their operational capabilities. Processing by Amedeo Ganciu.

Tab. 4. Overview table with analysis of nodes, breakdown of individual networks into communities and quality of clusters.

Network	Nodes	Edges	Community	Modularity
Authors	104'207	189'918	30'073	0.883
Keywords	95'942	569'779	1'859	0.612
Citations between articles	1'361'821	1'853'915	41'757	0.838

Mapping the networks through different algorithms can facilitate the understanding of the internal dynamics and relationships in each network. These tools were chosen considering their analytical and representational capacity according to the different characteristics of the networks: weighted/unweighted, oriented/unoriented, number of nodes, shortest possible response time (tab.5).

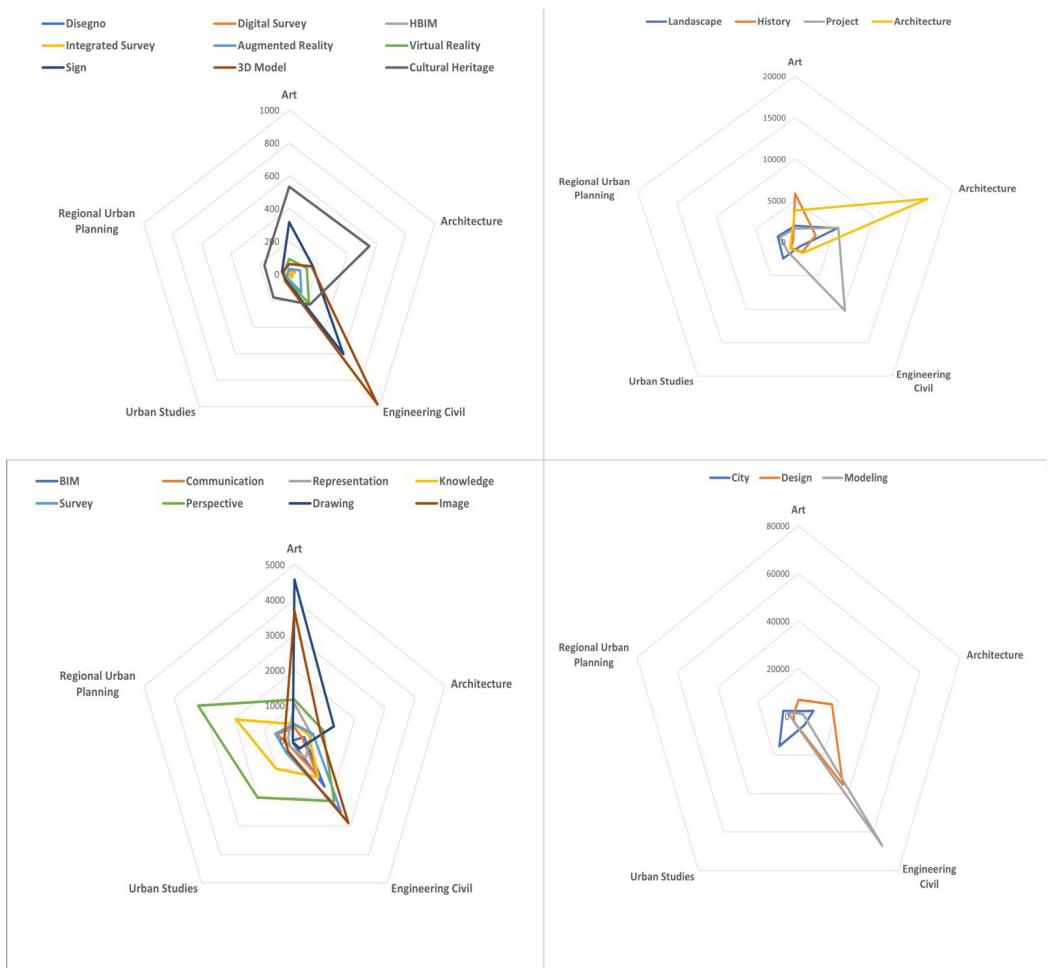



Fig. 04. Distribution of downloaded articles by WoS disciplinary categories. Processing by Amedeo Ganciu.

Tab. 5. Different layout algorithms for visualizing network layouts .

	Network		
Algorithm	Authors	Keywords	Citations between articles
OpenOrd	Yes	Yes	No
Force Atlas 2	Yes	Yes	No
Yifan Hu	No	No	Yes
Radial Axis	Yes	Yes	Yes

The visualisation of the author network confirms the numerical results for high complexity and fragmentation of collaborations between researchers, which can be observed with sufficient effectiveness by applying the OpenOrd partition layout and highlighting with different colours each of the more than 30,000 author families, i.e. the researchers who tend to collaborate with each other most frequently (fig. 05a).
The complementary ForceAtlas2 algorithm does not provide equally satisfactory graphical results for understanding the high fragmentation in these cases (fig. 05b). However, working with a cleaner network of authors, i.e. filtering it by considering those with the highest number of collaborations and obtaining a network of "only" 25'974 nodes and 80'006 connections the ForceAtlas2 layout develops a more interpretable graphical processing, allowing to observe the behaviour of the different families of Authors (fig. 06).
A closer inspection of the same layout reveals the collaborations between the top 130 authors within the top ten communities, i.e. those composed of the largest number of authors (fig. 07b). The Radial Axis ranking algorithm obviously allows us to discretise the relative importance of each author within each research group, but limits the understanding of the collaborations between the different research groups (fig. 07a). The particular result obtained could also be due to the evident tendency of the network towards engineering sectors, in particular those related to the development of new technological tools for representation, a sector in which the Asian area has been marking its future evolution for several years.
The implementation of the visualisation algorithms to the keyword network allows us to understand which are the hottest search topics and how they are connected to each other, obviously the different algorithms provided different results in terms of graphical comprehension. In this case, the complexity of the network is even greater in terms of the number of connections between the nodes, in fact, both the OpenOrd (fig. 08a) and the ForceAtlas2 (fig. 08b) did not allow appreciable elaborations, that is, they did not allow an effective interpretation of the structure. As in the previous case, a filter on the most frequent keywords makes it possible to reduce the complexity of the network in order to better appreciate the differences between the two layout tools with the OpenOrd algorithm (fig. 09b), which allows us to better emphasise the differences between the composition of the communities and the relations established between them with respect to ForceAtlas2 (fig. 09a).
By a further focus on the main keywords contained in the most populated communities, it is possible to identify some research themes that characterise the network. It can be observed that the topics of "Design" are strongly connected to the study of structures in different

Fig. 05. The Author Collaboration Network visualised with OpenOrd (A) and ForceAtlas2 (B), each node represents an author and its size is proportional to its research activity, arcs between nodes indicate collaboration between two authors, nodes of the same colour belong to the same community. Processing by Amedeo Ganciu.

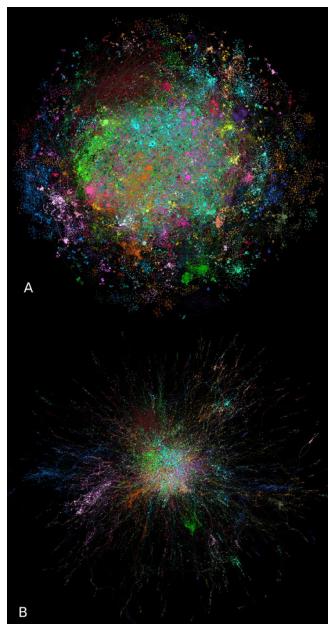
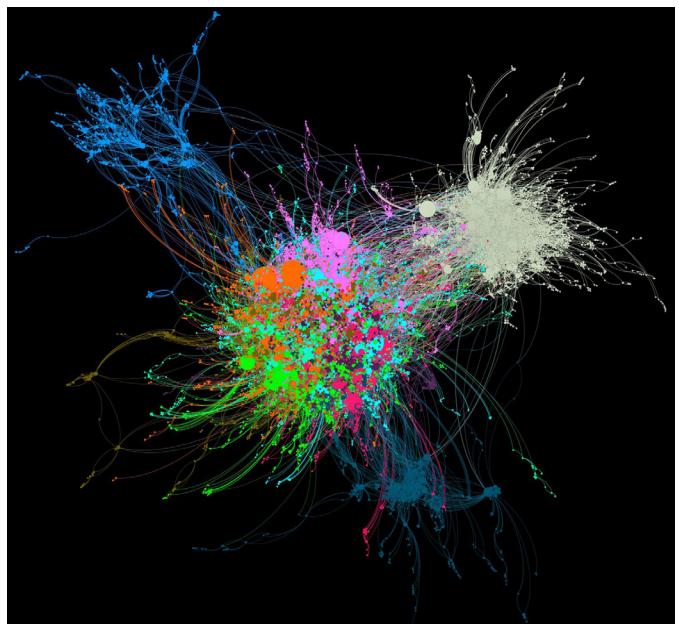



Fig. 06. The collaboration network between authors filtered and visualised with ForceAtlas2, each node represents an author and its size is proportional to its research activity, the arcs between nodes indicate collaboration between two authors, nodes of the same colour belong to the same community. Processing by Amedeo Ganciu.

aspects and through "Simulation" joins the other magenta community apparently dominated by "BIM" and "Project Management" (fig.10). In the third network, that of the citations, the high number of communities allows an effective visualisation of the fragmentation even when considered in its entirety, i.e. by displaying more than 1,000,000 nodes (fig. 11).
On further investigation using only the layout algorithms ForceAtlas2 (fig. 12a) and Yifan Hu (fig. 12b), it is possible to better observe the structure and connection between article families, i.e. groups of publications that tend to cite each other very frequently.

Conclusions

The research we presented was aimed at translating into images data and information that are difficult to analyse and interpret. In particular, the aim was to develop visualisations that would be useful for understanding the configuration and relationships that define the international scientific community around graphic sciences, starting from the keywords that define the main areas of research. The results make it possible to highlight some peculiarities useful for interpreting the relationships between the different researchers, the different topics of investigation and the different products of the scientific community. From the distribution of downloaded articles by *WoS* subject categories (fig. 05) one can highlight the associations between keywords and macro fields of knowledge. These

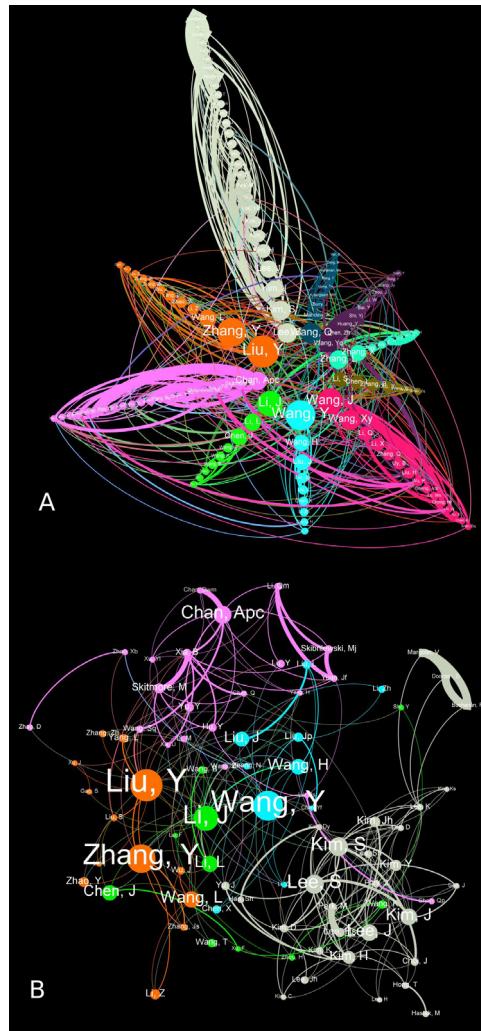



Fig. 07. The collaboration network between the main authors of the most important communities visualised with RadialAxis(A) and ForceAtlas2, each node represents an author and its size is proportional to its research activity, arcs between nodes indicate collaboration between two authors, nodes of the same colour belong to the same community. Processing by Amedeo Ganciu.

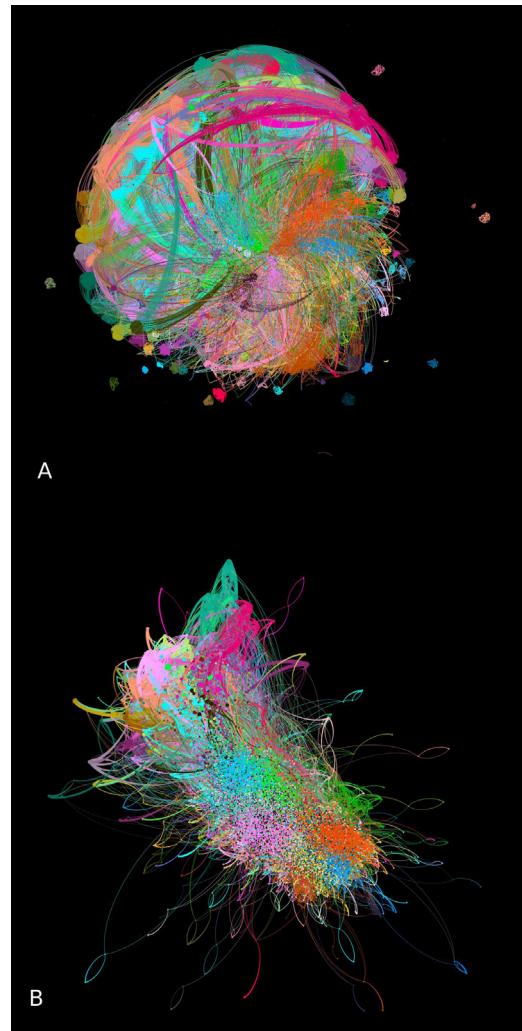



Fig. 08. Network of citations between keywords of downloaded publications displayed with OpenOrd (A) and ForceAtlas2 (B), each node represents a keyword of an article and its size is proportional to its frequency in the network, arcs between nodes indicate the co-presence of keywords on the same publication, nodes of the same colour belong to the same community. Processing by Amedeo Ganciu.

associations are also useful to place the research interests, summarised by the keywords, within the ERCs. For example, it can be seen that, in the international debate, research focused on 3D modelling falls within the strictly engineering sphere of interest, while research on drawing is more associated with the arts. In the visualisation of the network of collaboration between the main authors of the most important communities (fig. 08), there are clusters of relationships that can be interpreted as citation communities, which would merit further investigation in order to understand the bibliometric citation system. On the other hand, the visualisation of the keyword network (fig. 11) allows us to visualise

Figura 09. Keyword network filtered and visualised with ForceAtlas2 (A) and with OpenOrd (B), each node represents a keyword of an article and its size is proportional to its frequency in the network, the arcs between the nodes indicate the co-presence of the keywords on the same publication, the nodes of the same colour belong to the same community. Processing by Amedeo Ganciu.

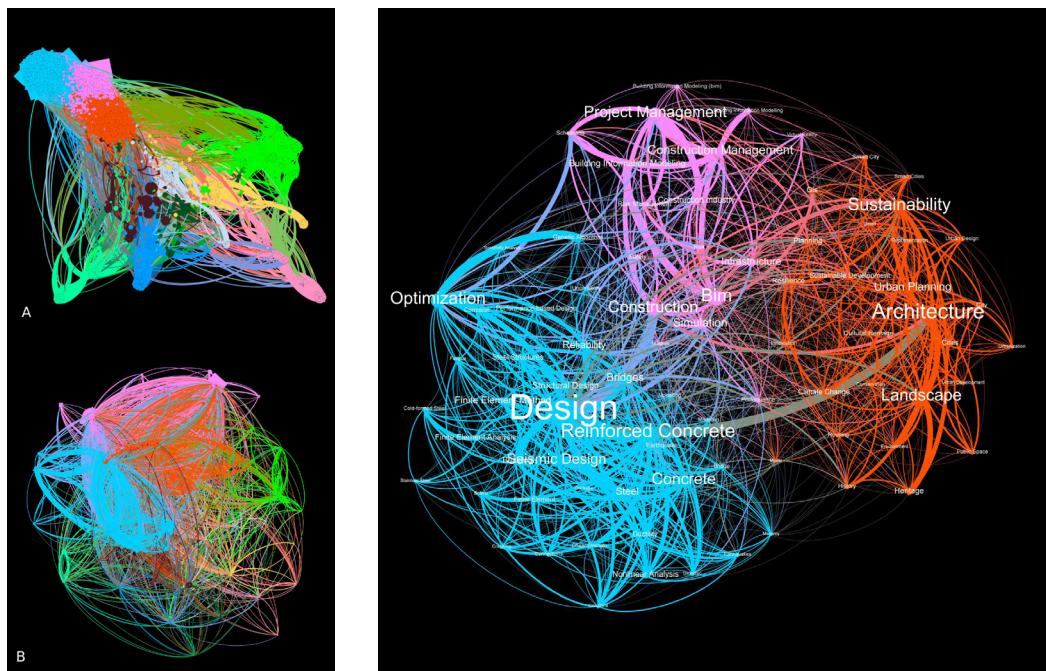
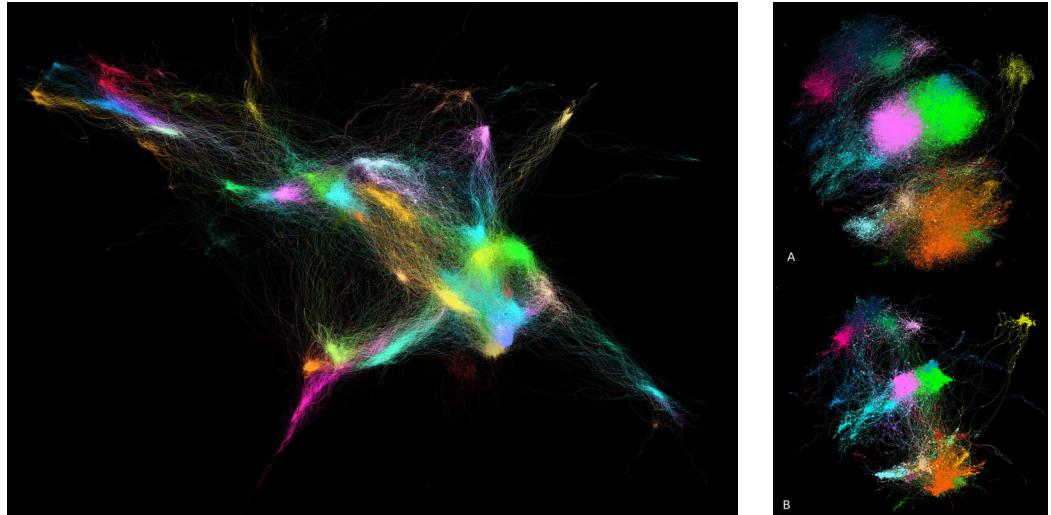
Fig. 10. Detail of the keywords network displayed with OpenOrd, each node represents a keyword of an article and its size is proportional to its frequency in the network, the arcs between the nodes indicate the co-presence of the keywords on the same publication, nodes of the same colour belong to the same community. Processing by Amedeo Ganciu.

the thematic densities within the scientific products and to define a nebula of research interests of the scientific community operating in the field of the disciplinary sectors defined by *WoS* within which the scientific community of the design can recognise itself. Finally, the application of Social Network Analysis to study the scientific landscape and its subsequent visualisation through various layout algorithms proves to be a valid and effective investigation tool.
The results of the analysis obtained from the massive download of publications from *WoS* are approximately in line with what can be hypothesized to be associated with the discipline of graphic sciences, however, it is worth remembering once again that for the Italian case most of the scientific production, not being catalogued through bibliometric structures, has escaped this research, therefore, in order to understand how the Italian "sector" collaborates with international production, in the next research it will be necessary to previously operate a meticulous work of gathering the information in order to then transform it into the ISI format. file format and then incorporate them within the structures illustrated in this research.

Fig. 11. Citation network between scientific articles displayed with ForceAtlas2 layout algorithm, each node represents a publication, the arcs between nodes indicate a citation relationship, given the complexity it is not possible to observe hierarchies between articles at this scale, nodes of the same colour belong to the same community. Processing by Amedeo Ganciu.

Fig. 12. Citation network between articles visualised with ForceAtlas2 (A) and Yifan Hu (B), each node represents a publication, its size is proportional to the number of citations received, arcs between nodes indicate a citation relationship, nodes of the same colour belong to the same community. Processing by Amedeo Ganciu.

References

Abt, H. A. (1998). Why some papers have long citation lifetimes. In *Nature*, n. 395(6704), pp.756-757.

Balestrieri, M., & Ganciu, A. (2019, July). The Use of Graphs to Explore the Network Paradigm in Urban and Territorial Studies. In *International and Interdisciplinary Conference on Image and Imagination*, pp. 1120-1132. Springer, Cham.

Batagelj, V. (2003). Efficient algorithms for citation network analysis. *arXiv preprint* cs/0309023.

Blondel, V. D. (2008). Fast unfolding of communities in large networks. In *Journal of statistical mechanics: theory and experiment*, 2008(10), P10008.

Boccaletti, S. (2006) Complex networks: Structure and dynamics. *Physics Reports*, No. 424, pp. 175–308.

Bocconcino, M., Vozzola, M. Infografica e visualizzazione grafica: nuovi modi per la rappresentazione dei dati. In *DIENNE V.7*, 7/2020.

Börner, K., Chen, C., Boyack, K. W. (2003). Visualizing knowledge domains. In *Annual review of information science and technology*, 37(1), 179-255.

Caldarelli, G. (2007). Scale-free networks: complex webs in nature and technology. In *Oxford University Press*.

Cicalò Enrico (2020a). Connessioni tra saperi: disciplinarietà, interdisciplinarità e transdisciplinarietà delle scienze grafiche/ Connections between knowledge: disciplinarity, interdisciplinarity and transdisciplinarity of graphic sciences. In Arena A., "et al." (a cura di). *Connettere. Un disegno per annodare e tessere.* Atti *del 42° Convegno In- ternazionale dei Docenti delle Discipline della Rappresentazione/Connecting. Drawing for weaving relationships. Proceedings of the 42th International Conference of Representation Disciplines Teachers*. Milano: FrancoAngeli, pp. 316-337.

Cicalò, E. (2020b). Exploring Graphic Sciences. In *International and Interdisciplinary Conference on Image and Imagination* (pp. 3-14). Springer, Cham.

Cicalò, E., Menchetelli, V., (2020) Informazioni accessibili. la visualizzazione dati nell'epoca dei big data, degli open data e degli open tools. In *DIENNE V.7*, 7/2020.

Cicalò, E., Pileri, M., Valentino, M., (2021). Connessione tra saperi. Il contributo delle scienze grafiche nella ricerca in ambito medico. Connecting Knowledge. The Contribution of Graphic Sciences to Medical Research. In A. Arena, M. Arena, D. Mediati, P. Raffa, *42° Convegno Internazionale dei Docenti delle Discipline della Rappresentazione CONNETTE-RE_un disegno per annodare e tessere Linguaggi, Distanze, Tecnologie*. 16, 17 e 18 settembre 2021, pp. 510-527. Reggio Calabria e Messina.

Cook, E.A. (2002) Landscape structure indices for assessing urban ecological networks. In *Landscape. Urban Planning.* 58, pp.269–280.

De Meo (2011, November). Generalized louvain method for community detection in large networks. *In 11th international conference on intelligent systems design and applications*, pp. 88-93. IEEE.

Derrible, S.; Keendy, C. (2011) Applications of Graph Theory and Network Science to Transit Network Design. In *J. Transp. Rev.*, No. 31, pp. 495–519.

Estrada, E., Bodin, Ö. (2008) Using network centrality measures to manage landscape connectivity. *In Ecol. Appl.*, No. 18, pp. 1810–1825.

Ganciu, A.; Balestrieri, M.; Cicalò, E. (2016, June) Visualising the research on visual landscapes. Graph representation and network analysis of international bibliography on landscape. In *Proceedings of the XIV International Forum Le Vie dei Mercanti*, Capri, Italy, 16–18 June 2016.

Ganciu, A. (2018). Dynamics of metropolitan landscapes and daily mobility flows in the Italian context. An analysis based on the theory of graphs. In *Sustainability*, Vol. 10(3), p. 596.

Garfield, E., Sher, I. H., Torpie, R. J. (1964). The use of citation data in writing the history of science. Institute for Scientific Information Inc Philadelphia PA.

Gigliarelli, E. (2020) La rappresentazione della conoscenza utilizzando le mappe cognitive per favorire la ricerca collaborative. In *Building Information Modeling, Data & Semantics*, 7/2020.

Green, N. (2007). Functional polycentricity: A formal definition in terms of social network analysis. *In Urban Stud.*, No. 44, pp.2077–2103.

Hu, Y.F., (2005) Efficient and high quality force-directed graph drawing. In *The Mathematica Journal*, 10 (37-71), 2005.

Jin, J. H., Park, S. C., Pyon, C. U. (2011). Finding research trend of convergence technology based on Korean R&D network. In *Expert Systems with Applications*, 38(12), 15159-15171.

Lo, H.K.; Tung, Y.K. (2003) Network with degradable links: Capacity analysis and design. In *Transp. Res. Part B Methodol.* 37, pp. 345–363.

Luigini, A. (2021). Ricerca interdisciplinare e ICAR17: una proposta per la definizione di un modello condiviso. *42° Convegno Internazionale dei Docenti delle Discipline della Rappresentazione CONNETTERE_un disegno per annodare e tessere Linguaggi, Distanze, Tecnologie*. 16, 17 e 18 settembre 2021, Reggio Calabria e Messina.

Luigini A (2021b); Riviste scientifiche nel settore ICAR17: analisi quantitativa delle keywords e dei temi di ricercar. In A. Arena, M. Arena, D. Mediati, P. Raffa, *42° Convegno Internazionale dei Docenti delle Discipline della Rappresentazione CONNETTERE_un disegno per annodare e tessere Linguaggi, Distanze, Tecnologie*. 16, 17 e 18 settembre 2021, pp. 510-527. Reggio Calabria e Messina.

Martin S. (2011) OpenOrd: An Open-Source Toolbox for Large Graph Layout. In *SPIE Conference on Visualization and Data Analysis (VDA)*.

Mathis, P. (2007). Graphs and Networks: Multilevel Modeling.

Newman, M. E. (2003). The structure and function of complex networks. SIAM review, 45(2), pp. 167-256.

Newman, M.E.J.; Girvan, M. (2004) Finding and evaluating community structure in networks. In *Phys. Rev. E.* 69, pp. 1–15.

Patuelli, R. (2007) Network analysis of commuting flows: A comparative static approach to German data. In *Netw. Spat. Econ.* No.7, pp. 315–331.

Porta, S.; Crucitti, P.; Latora, V. (2006). The network analysis of urban streets: A dual approach. Phys. A Stat. Mech. Its Appl. No. 369, pp. 853–866.

Reuters, T. (2016). Web of knowledge-real facts-ip & science-thomson reuters.

Robins, G. (2013). A tutorial on methods for the modelling and analysis of social network data. In J. Math. *Psychol.* No. 57, pp 261–274.

Salas-Olmedo, M.H.; Nogués, S. (2012). Analysis of commuting needs using graph theory and census data: A comparison between two medium-sized cities in the UK. In *Appl. Geogr.* No. 35, pp. 132–141.

Urban, D.L.; Minor, E.S.; Treml, E.A.; Schick, R.S. (2009) Graph models of habitat mosaics. In Ecol. Lett. No. 12, pp.260–273.

Xie, F.; Levinson, D. Evolving (2011) Transportation Networks. Berlin:Springer Science & Business Media.

Authors
Amedeo Ganciu, Dipartimento di Architettura, Design e Urbanistica, Università degli Studi di Sassari, aganciu@uniss.it
Andrea Sias, Dipartimento di Architettura, Design e Urbanistica, Università degli Studi di Sassari, a.sias19@studenti.uniss.it

To cite this chapter: Ganciu Amedeo, Sias Andrea (2022). Visualizzare la conoscenza. La rappresentazione delle reti citazionali internazionali nell'ambito delle scienze grafiche/Visualising the knowledge. The representation of international citation networks in the graphic sciences. In Battini C., Bistagnino E. (a cura di). *Dialoghi. Visioni e visualità. Testimoniare Comunicare Sperimentare. Atti del 43° Convegno Internazionale dei Docenti delle Discipline della Rappresentazione/Dialogues. Visions and visuality. Witnessing Communicating Experimenting. Proceedings of the 43rd International Conference of Representation Disciplines Teachers*. Milano: FrancoAngeli, pp. 2479-2502.

 Isbn 9788835141938